PNL24

THE ESTIMATION POWER OF ALTERNATIVE COMORBIDITY INDEXES

Baser O[1], Stephenson J[2]
[1]Thomson-Medstat, Ann Arbor, MI, USA, [2]Thomson Medstat, Philadelphia, PA, USA

OBJECTIVES: Health care expenditures are strongly influenced by overall illness burden. Appropriate risk adjustment is required for correct policy analysis. We compared three risk adjustment methods—Charlson Comorbidity Index (CCI), Elixhauser (ELX), and Chronic Disease Score (CDS)—in terms of their estimation power in analyses of healthcare expenditures. **METHODS:** Seven models were considered. Using the same demographic factors, models were separated by index variables: 1) CHS only; 2) ELX only; 3) CDS only; 4) CHS and ELX; 5) CHS and CDS; 6) CDS and ELX; and 7) CHS, ELX, and CDS. A generalized linear model with log link and gamma family was used to estimate the models. BIC, AIC, and log likelihood scores were calculated across the models to see which model afforded the best fit. Average squared prediction error (ASPE) was considered to asses the estimation power of these indexes. **RESULTS:** MarketScan® data were used to estimate the total health care expenditures of migraine patients treated with a triptan. After applying inclusion and exclusion criteria, we identified 43,776 migraine patients who used a triptan and we used this population to create the analytic samples. CCI, an older and common risk adjustment method, performed the poorest in terms of estimation power. Of the single-index models (models 1–3), ELX performed the best; of the double-index models (models 4–7), CCI and ELX performed the best. We conducted a detailed analysis of multicollinearity using a variance inflation factor and failed to find any multicollinearity among the three measures. Overall, model 7, where we used three indexes in the same model, performed the best. **CONCLUSION:** We found that the three risk adjustment indexes measure different risks and that the indexes can be used together in a single model. Using only CCI is misleading.

PNL25

FAILURE OF REGRESSION ADJUSTMENT AGAINST PROPENSITY SCORE MATCHING

Baser O
Thomson-Medstat, Ann Arbor, MI, USA

OBJECTIVES: Causal inference is challenging in observational studies because of inevitable presence of self-selection: treatment group are usually different from control group in terms of risk factors. Regression adjustment and propensity score matching are commonly used methods to adjust for confounders. In this paper, we show when regression adjustment fails to adjust for differences in observed covariates and propensity score matching is the only alternative. **METHODS:** The following guidelines are provided to decide which method should be used. Multivariate analysis fails if 1) the distributions of the covariates in both groups are not symmetric; 2) the sample sizes are different; 3) the distributions of the covariates in both groups have different variance and the means of propensity scores in the two groups are more than half a standard deviation apart; 4) the ratio of the variances of the propensity score in the two groups differs from one; or 5) the ratio of the variances after residuals of the covariates (including propensity score) is different from one. **RESULTS:** To apply the suggested guidelines, we conducted a retrospective cohort study analyzing the impact of triptan use on total healthcare expenditures among patients with migraine. Medstat MarketScan® data were used for the analysis. We applied the suggested guidelines, and found that propensity score matching, rather than multivariate analysis, should be used to estimate treatment effect. The means of propensity score were two standard deviations apart, the Smirnow test showed that the distributions were not symmetric ($p = 0.0000$), sample sizes were different (treatment = 43,799 vs. control = 128,366), and bootstrapping results showed that the ratio of variance of the propensity score in the two groups and in residuals were significantly different from one. **CONCLUSIONS:** Results from regression analysis can be misleading and propensity score matching should be used as an alternative under certain conditions.

PNL26

EXTENDING MATCHING ESTIMATORS OF CAUSAL EFFECTS TO CONSIDER UNOBSERVED VARIABLE BIAS: AN APPLICATION OF SENSITIVITY ANALYSIS

Devine JW, Hadsall RS, Farley JF
University of Minnesota, Minneapolis, MN, USA

OBJECTIVES: Unobserved variable bias remains a major limitation of observational studies that rely on matching estimators to examine causal effects. Several methods have been proposed to quantify the extent of uncertainty about unobserved variable bias. The primary aim of this study was to illustrate Rosenbaum's method of sensitivity analysis to examine how inferences from observational research may change in the presence of varying amounts of unobserved variable bias. **METHODS:** Data used in the illustration were collected from an administrative claims database as part of an observational study that evaluated the effect of daily migraine prevention on the consumption of migraine-specific abortive medication. The database included information on all prescription services rendered between 1 October 2002 and 30 September 2004 for a population of migaineurs in the Military Health System. Each patient receiving treatment (daily migraine prevention) during the observation period was matched to a similar untreated patient via nearest neighbor matching on a propensity score. The illustration of the sensitivity analysis was based on the Wilcoxon's signed rank statistic and attempted to estimate the magnitude of departure from equal treatment probabilities between matched pairs sufficient to alter the statistical conclusions of the study results. **RESULTS:** The analysis of 997 matched pairs suggested that treatment with daily migraine prevention was associated with a decline in the use of migraine-specific abortive medication. However, the sensitivity analysis indicated that the treatment effects were sensitive to assumptions about unobserved variable bias. Specifically, the probability of obtaining the results under the null hypothesis of no effect approached the conventional 0.05 level when the odds of receiving treatment within matched pairs differed by a factor of 1.4 or more. **CONCLUSIONS:** Rosenbaum's sensitivity analysis is a useful mechanism for discussing the magnitude of uncertainty surrounding unobserved variable bias in observational studies.

NEUROLOGICAL DISORDERS—Patient Reported Outcomes

PNL27

THE EFFECT OF ADHERENCE TO ALZHEIMER'S DISEASE TREATMENT ON HEALTH CARE COSTS IN MANAGED CARE

Harada ASM, Vanderplas AM
Prescription Solutions, Irvine, CA, USA

OBJECTIVE: To determine the effect of Alzheimer's disease (AD) treatment adherence on the change in overall health care costs following AD treatment initiation in managed care. **METHODS:** A retrospective analysis was conducted using pharmacy, medical, and member eligibility data from a large managed

care organization. Members initiating therapy between January 1, 2004 and June 30, 2004 with one of the following five AD regimens were included in the analysis: donepezil, galantamine, rivastigmine, memantine, or memantine+AChEI (acetylcholinesterase inhibitor) combination. We excluded members with previous AChEI use or who were not continuously enrolled throughout the 360-day pre-treatment and 360-day post-treatment initiation periods. Adherence was defined as the total days supplied (maximum of 360) of the AD medication initiated divided by the total days in the post-period (360 days). Change (post-treatment minus pre-treatment) in pharmacy, medical, and total (pharmacy+medical) costs was analyzed, adjusting for age, gender, comorbidities (chronic disease score), AD regimen, and pre-treatment total costs. **RESULTS:** A total of 2405 members were newly started on AD therapy; 65% were female and mean age was 81 years (SD 7.1). Treatment was initiated with donepezil (62%), memantine (15%), galantamine (11%), rivastigmine (8%), and memantine+AChEI (6%). Mean adherence to AD therapy was 0.57 (SD 0.36). The adjusted analysis revealed that a 0.1 adherence rate increase (for example, an increase in adherence from 0.5 to 0.6) was associated with an increase of \$215 ($p < 0.001$) in annual pharmacy costs and a decrease of \$783 ($p < 0.001$) and \$568 ($p = 0.008$) in annual medical and total health care costs, respectively. **CONCLUSIONS:** An increase in AD therapy adherence is associated with overall health care cost savings in the first year following treatment initiation.

PNL28

STUDYING DISCONTINUATION, SWITCHING, AND AUGMENTATION USING COMPETING RISK METHODS

Gause D, Arcona S

Novartis Pharmaceuticals, East Hanover, NJ, USA

OBJECTIVES: To illustrate use of competing risk methodology for studying risk of discontinuing, switching, or augmenting medications used to treat bipolar disorders. **METHODS:** Administrative claims data (MarketScan®, Medstat) was used to identify patients with bipolar disorders treated with montherapy atypical antipsychotics. Patients were followed 12 months to identify those augmenting (starting second medication class within 30 days and before any discontinuation), discontinuing (medications runs out and no refill within 30 days following expiration of days supply), and switching (patient discontinues medications and starts different class within 30 days). Because risks of augmentation, discontinuation, and switching are not independent, competing risk methods are used to summarize likelihood of different outcomes. Cumulative incidence functions are computed as a function of crude hazards for each of the 3 dependent outcomes. Plots of the cumulative incidence functions provide estimates of the probability of each outcome occurring by any time point. **RESULTS:** A total of 787 patients on monotherapy atypical antipsychotics were identified. By 6 months 21% of the patients had discontinued, 3% had switched, 56% had augmented, and the remaining 20% remained on their initial antipsychotic. By one year 29% of the patients had discontinued, 4% had switched, 59% had augmented, and 8% remained unchanged. **CONCLUSIONS:** Switching and augmentation are outcomes important to study along with discontinuation. Treating these as independent when estimating risk of each (e.g. with complement of Kaplan-Meier estimates) would give biased estimates of the marginal probabilities. Cumulative incidence curves provide a means of dynamically seeing the risks of all three outcomes.

PNL29

A REVIEW OF THE VALIDITY AND RELIABILITY OF THE PARKINSON'S DISEASE QUESTIONNAIRE (PDQ-39)

Stafkey-Mailey DR[1], Gwadry-Sridhar F[2]

[1]University of Southern California, Los Angeles, CA, USA, [2]University of Western Ontario, and London Health Sciences Centre, London, ON, Canada

OBJECTIVE: Patients with Parkinson's Disease (PD) have diminished health related quality of life (HRQoL) due to a number of emotional, functional and physical limitations. We examined the reliability and validity of the PDQ-39 as both a discriminant and evaluative instrument for use in Parkinson's disease studies. **METHODS:** We conducted a literature review using MEDLINE and HealthStar citing the PDQ-39, from 1995–2005 using the terms "Parkinson's Disease Questionnaire or PDQ-39" limited to Humans and English language. We reviewed each publication for evidence supporting the instrument's reliability face and content validity, construct validity measured as the correlation of the PDQ-39 responses with other relevant instruments and clinical measures, and test-retest reliability or intra class correlation coefficients. Discriminant abilities were evaluated by looking for evidence of significance between PD patients with varying levels of disability. Responsiveness was evaluated using change scores and established minimally important difference (MID). **RESULTS:** Thirty-nine papers met the inclusion criteria. Correlation of scales of the PDQ-39 with the Medical Outcomes Study 36-item Short Form (SF-36) showed strong correlation (>0.5). Intraclass correlation coefficients for each domain of the PDQ-39 ranged from 0.67 to 0.96 across studies. Mean scores for each domain were shown to be significantly statistically different ($p < 0.01$) between PD patients with Hoehn & Yahr rating I, II, III, and IV. Changes in PDQ-39 scores were significantly correlated with changes in patients' retrospective judgments of change (0.25–0.31, $p < 0.01$), and changes on the SF-36 (0.25–0.37, $p < 0.05$), but not with clinical assessments. Two studies calculated a MID for the PDQ-39 which varied between 5–10 points on the summary index and 2–11 points across domains. **CONCLUSION:** There is substantial evidence supporting the reliability and validity of the PDQ-39 as a discriminate instrument, but further research on its usefulness as an evaluative instrument is warranted to increase our confidence.

PNL30

COMPARISON OF HEALTH-RELATED QUALITY OF LIFE QUESTIONNAIRES IN MULTIPLE SCLEROSIS: A REVIEW

Pawar VS, Miller LA

West Virginia University, Morgantown, WV, USA

Multiple sclerosis (MS) being a progressive neurological disease with no cure, lends itself as an important area for health-related quality of life (HRQOL) research. A number of questionnaires have been designed to measure HRQOL in MS patients, but no one measure has been established as a gold standard. The large number of questionnaires also gives neurologists a range of choices but limited information on which to base their selection. In order to determine the relative advantages and disadvantages of the HRQOL questionnaires, a comparison of their psychometric properties in the same patient population is necessary. **OBJECTIVE:** To review the current literature and identify studies comparing psychometric properties of MS-related QOL measures within the same patient population. **METHODS:** Studies included were those that administered more than one HRQOL questionnaire in the same group of patients with MS. The search was restricted to the PUBMED and MEDLINE databases for articles published between 1991 and 2005. **RESULTS:** Of almost 30 MS-related QOL questionnaires, psychometric